SECURITY CLASSIFICATION OF THIS PAGE

PORT DOCUMENTATION PAGE

AD-A211 138

1b RESTRICTIVE MARKINGS

2a

2b

3 DISTRIBUTION / AVAILABILITY OF REPORT
Approved for public release; distribution unlimited.

4 PERFORMING ORGANIZATION REPORT NUMBER(S)

5 MONITORING ORGANIZATION REPORT NUMBER(S)
AFOSR-TR- 89-1057

6a NAME OF PERFORMING ORGANIZATION
New York Univeristy

6b OFFICE SYMBOL (If applicable)

7a NAME OF MONITORING ORGANIZATION
Air Force Office of Scientific Research/NL

6c ADDRESS (City, State, and ZIP Code)
Psychology Department
New York, NY 10003

7b ADDRESS (City, State, and ZIP Code)
Building 410
Bolling AFB, DC 20332-6448

8a NAME OF FUNDING / SPONSORING ORGANIZATION
AFOSR

8b OFFICE SYMBOL (If applicable)
NL

9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
AFOSR-88-0140

8c ADDRESS (City, State, and ZIP Code)
Building 410
Bolling AFB, DC 20332-6448

10 SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO
61102F	2313	A5	

11 TITLE (Include Security Classification)
Ratings of Kinetic Depth in Multi-Dot Displays

12 PERSONAL AUTHOR(S)
George Sperling

13a TYPE OF REPORT	13b TIME COVERED FROM ____ TO ____	14 DATE OF REPORT (Year, Month, Day)	15 PAGE COUNT
Reprint			43

16 SUPPLEMENTARY NOTATION

17 COSATI CODES		
FIELD	GROUP	SUB-GROUP

18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

DTIC
ELECTE
AUG 11 1989
S D
D

20 DISTRIBUTION / AVAILABILITY OF ABSTRACT
☒ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT ☐ DTIC USERS

21 ABSTRACT SECURITY CLASSIFICATION
UNCLASSIFIED

22a NAME OF RESPONSIBLE INDIVIDUAL
Dr. John F. Tangney

22b TELEPHONE (Include Area Code)
(202) 767-5021

22c OFFICE SYMBOL
NL

DD FORM 1473, 84 MAR

83 APR edition may be used until exhausted.
All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

89 8 10 061

AFOSR-TR- 89-1057

Mathematical Studies in Perception and Cognition

87 - 3 (Rev. 2)

Ratings of Kinetic Depth in Multi-Dot Displays

Barbara A. Dosher,[*] *Michael S. Landy,*[**] *and George Sperling*[**]

* Psychology Department
Columbia University

** Psychology Department
New York University

Running Head: KDE Ratings

Address correspondence to:

Barbara A. Dosher
Psychology Department
New York University
6 Washington Place, Rm. 980
New York, New York 10003

Accesion For
NTIS CRA&I
DTIC TAB
Unannounced
Justification
By
Distribution /
Availability Codes
Dist | Avail and / or Special
A-1

Ratings of Kinetic Depth in Multi-Dot Displays

Barbara A. Dosher,[*] *Michael S. Landy,*[**] *and George Sperling*[**]

ABSTRACT

Subjects viewed kinetic depth displays whose shape (sphere or cylinder) was defined by luminous dots distributed randomly on the surface or throughout the volume of the object. Subjects rated the amount of perceived 3D depth, rigidity, and coherence. (Coherence is high when all the dots are perceived as a single object.) There were significant individual variations in ratings but, on the whole, all three ratings increased with the number of dots. Luminance of dots had no effect on any of the ratings. Points within the volume yielded ratings equal to or greater than surface points. Each of the three ratings varied with at least 3 of the 4 factors (shape. distribution, numerosity, and perspective), but the ratings did not necessarily covary either between trials or between conditions–often they were uncorrelated or negatively correlated. For example, object shape affected ratings of rigidity but not of depth; when perceived veridically, polar perspective displays were rated slightly less rigid than parallel projection displays but they received higher depth ratings. (When perceived in reversed perspective, polar displays were grossly nonrigid, independent of the other factors.) The complex but understandable interplay of stimulus parameters and ratings forces a re-examination of previous experimental results and theories in which different KDE ratings were treated interchangeably.

When a two-dimensional (2D) projected image corresponds to a three-dimensional (3D) object that is rotating, viewers frequently perceive an object with depth. Because rotation induces apparent three-dimensional depth even when isolated still views of the object fail to induce perceived depth, the phenomenon is called the *kinetic depth effect* or KDE (Wallach and O'Connell, 1953). In this paper, experiments are discussed which consider the perception of dot displays, where each stimulus consists of illuminated dots on an otherwise invisible object. It will be demonstrated that there are a number of partially de-coupled aspects to the perception of these displays under motion: (1) Coherence – whether all dots in the display are seen as constituting a single object. (2) Depth – the amount of 3D depth seen in the display. (3) Rigidity – whether those illuminated dots that are perceived as constituting a coherent object also are perceived as maintaining their relative 3D positions (rigid appearance), or as changing their relative positions (nonrigid, rubbery appearance).

There is a large body of literature examining the function of various kinds of stimulus variables in the kinetic depth effect. Some of the classic stimulus variables include the number of elements defining the stimulus, element shape, occlusion, perspective, correspondence, element density, and rotation speed. The effects of these stimulus variables are examined by a variety of dependent measures: global "goodness" judgments (Andersen & Braunstein, 1983; Braunstein, 1962; Braunstein & Andersen, 1984; Green, 1961; Petersik, 1980); qualitative motion categorization (surface, rotary, oscillatory – Caelli, 1979, 1980); judgments about objective rotation direction (Braunstein, 1962, 1977; Petersik, 1979, 1980); perceived curvature or shape (Braunstein & Andersen, 1984; Todd, 1984); or proportion of corresponding elements

across frames (Lappin, et. al., 1980). One question which arises is whether the choice of dependent measure is of no consequence: are all these measures simply reflections of a unitary aspect of the kinetic depth percept? In order to answer this question, we examined the independence of the three aspects of the percept listed above by collecting three separate responses on every trial in experiments which varied some important stimulus variables.

To be more concrete, consider an early set of experiments by Green (1961). He examined, among other factors, the importance of number of stimulus elements on the KDE. Subjects were asked to rate displays on a scale which combined the notions of rigidity and coherence, as defined here. The label "goodness" is used here to describe Green's combined rating scale, in order to distinguish it from our use of the distinct labels rigidity and coherence (which Green used interchangeably). Green demonstrated that the number of stimulus elements was a potent factor in determining the "goodness" of a perceived object under various forms of rotation; generally, the more stimulus elements, the higher the rated "goodness", with the largest increments occurring with numbers of elements under 32. In principle, the increment in "goodness" could have reflected some unspecified weighting of coherence and rigidity. It is not clear whether numerosity affects one or both of these aspects of the percept primarily, nor is it clear how it affected perceived depth of element trajectories.

Here we investigate element numerosity, as well as a number of other factors that may vary in viewing 2D projected images of objects. In particular, we examine one image projection factor: parallel projection versus perspective projection (projection distance at 3

times object radius); and three object factors: number of elements representing an object (from 4 elements to 80 elements), whether the elements representing the object were entirely on the surface, or distributed throughout the volume, and the strength of density cues to depth in a still frame (by using different forms). Our aim was to determine the effect of these KDE stimulus variables on ratings of coherence, depth, and rigidity.

Our results will corroborate some findings of previous investigators, for example, that the number of dots in the object and the presence of polar perspective can add to the strength of KDE. However, we will show that these stimulus variables do not generally affect all three aspects of the KDE percept equally, and that there are many subtleties and complexities in the KDE.

Method

Because of the large number of stimulus variables, the study was divided into three separate experiments. The experiments were conducted with the same subjects and with the same procedures, except as noted.

Subjects.

There were four subjects in the experiments, including two of the authors and two students. Three subjects had normal or corrected-to-normal vision, subject CFS could be corrected only to 20:40.

Apparatus.

All stimuli were computer generated and the display and response collection was computer controlled. Experiment 1 and a pilot experiment used a point/vector display controller (Kropfl, 1975) and an HP1304A display monitor. Display resolution was 1024 x 1024. Experiments 2 and 3 used a raster display controller, Adage RDS-3000, and a Conrac 7211C19 RGB color monitor. Display resolution was 512 x 512. Experiment 1 used binocular viewing in a completely dark room. In Experiments 2 and 3, subjects viewed the stimuli monocularly through a reduction tube, with an aperture slightly larger than the stimuli. Hence, weak stereo cues to flatness may have been present in Experiment 1, but not in Experiments 2 and 3.

Stimuli.

Stimuli consisted of random white dots scattered on the surface or throughout the volume of invisible spheres and cylinders. The probability distribution used for dot placement was uniform across the surface (or through the volume) in each case, but choices of dots were constrained so as to fill the surface or volume fairly evenly by partitioning into equal area (or equal volume) segments and putting equal numbers of dots in each segment. Five stimulus parameters were varied. First, there were two types of objects, a sphere of diameter 2 *dva* (degrees of visual angle), and an upright cylinder of height 2 *dva* and cross-sectional diameter 2 *dva*. The number of dots was varied from 4 to 80. These dots were either positioned on the surface or in the volume of the object being simulated. Stimuli were either presented in

parallel projection (i.e. with no perspective), or with an exaggerated amount of polar perspective (corresponding to a viewing distance of 3 times the object radius, far smaller than the actual viewing distance). All stimuli were rotated about a vertical axis through the center of the simulated object. Stimuli were either rotating front-left or front-right, although this distinction is only meaningful for the stimuli with polar perspective. Single-frame views of some sample stimuli are shown in Figure 1.

Procedure.

On each trial, the subject was shown a fixation target, which then disappeared, and was followed shortly thereafter by one rotation of the stimulus. See Table 1 for details of rotation speeds, etc. After the stimulus presentation was complete (approximately 2 sec), four responses were required of the subject. First, the subject indicated the direction of rotation of the object (front-left or front-right). These responses were used in polar projection displays to determine whether the the subject perceived the object in the veridical or the reversed mode. Then, three different ratings of the percept were required: depth, coherence, and rigidity.

Depth rating. The subject indicated the amount of depth perceived in the stimulus on a scale from 1 to 5. Given that all stimuli were based upon objects rotating about a vertical axis, depth was related to an inferred "top view" of the stimulus. The subject was shown the top views (Figure 2) to facilitate his rating. The most depth, 5, was associated with a perceived circular path for each dot; the least depth, 1, was associated with no perceived depth, and

hence an oscillatory linear path for each dot.

Figures 1 and 2 about here.

Coherence rating. The next rating, also on a scale of 1 to 5, was of the perceived coherence of the multi-dot display. A rating of 5 indicates the greatest coherence (i.e. all the dots held together as one object). A rating of 4 indicates that a few dots did not cohere; 3 indicates that the display broke up into 2 distinct objects (segmentation); 2 indicates that three or more objects were perceived; 1 indicates there was no perceived coherence whatsoever.

Rigidity rating. Perceived rigidity was rated on a scale from 1 to 5, with a 5 indicating one or more totally rigid objects, and lower numbers indicating more and more non-rigidity or "rubberiness".

Subjects judge all three aspects of each percept. This allows us to relate the three aspects on an individual trial basis. Had the three judgments been collected separately, the relation between the three judgments would have been available only at the level of the mean data.

Designs.

Experiments 1, 2, and 3 differed in the factors varied and in the display devices used. Table 1 summarizes the design and viewing conditions for each experiment. All designs were fully crossed in each included factor. Left and right veridical rotation direction was also a

factor in each experiment. Stimulus order was randomized within block and each block consisted of one token of each stimulus type, yielding 64 stimuli per block in Experiment 1 and 32 stimuli per block in Experiments 2 and 3. A different random token of each stimulus type was generated for each of 6 blocks per experiment. A pilot experiment yielded no effect on any response ratings of object size (2 *dva* vs 4 *dva*); 2 *dva* was used subsequently. The polar perspective manipulation was defined with respect to object radii, so the object size manipulation also varied the mismatch between actual and appropriate viewing distance for the degree of perspective. This and some of the current work was originally reported in Landy, Dosher, and Sperling (1986).

Table 1 about here.

Results

The results for Experiments 1 and 2, pooled across subjects, are shown in Figure 3a and 3b. There were significant individual differences, discussed below, and so statistical analyses were performed as within-subject ANOVAs. The twelve trial-type replications that result from collapsing over rotation direction and test block formed the random factor. These replications represented responses to twelve distinct, randomly generated stimuli of each type. Table 2 lists the significance levels associated with each rating, factor, subject and experiment, along with a qualitative coding of the direction of the result. Table 2 thus gives a quick summary of the

consistency both between subjects and within subjects across experiments. Table 3 summarizes the results of previous related experiments. Notice that the current set of experiments include factors and ratings which are either unrepresented in the literature or represented by a questionable combined measure.

Figure 3 and Tables 2 & 3 about here.

Numerosity. When the number of dots ranges from very small to moderate in number (Expt. 1), increasing the number of dots increases all three ratings for three subjects of the four. The fourth subject shows a very different behavior (e.g. see Figure 4 below). The 4 dot stimuli yielded very high depth ratings for this subject. The subject mentioned afterward that the stimuli reminded him of organic chemistry drawings, and appeared very vivid.

When the number of dots was moderate to large (Expt. 2), the effect of number of dots was less dramatic. Depth ratings increased slightly and saturated at these high numerosities. At these larger levels of numerosity, the effects of numerosity on coherence and rigidity were small: there were no significant effects on coherence ratings and numerosity was related to rigidity ratings for only two of four subjects. In summary, increasing dot numerosity increases all three ratings, up to a point, and then saturates. Depth ratings appear to increase and saturate in a continuous manner, while coherence and rigidity ratings are high for all but sparse displays (8 or fewer elements). These findings are in general agreement with those of Green

(1961) over similar ranges of numerosity, however Green's judgment was one of overall "goodness", and more nearly agrees with the depth judgments here.

Intensity. When intensity of displays was increased from 0.86 to 42.7 μcd/dot (Expt. 3), there was no significant effect on ratings (with the exception of a single subject on a single rating). We rule out an effect of varying display types in which overall stimulus intensity (contrast) varies in the visible range. (But see Dosher, Landy, & Sperling, 1988, for manipulations of intensity very near to threshold, which do affect kinetic depth performance.)

Form. When a spherical shape was directly contrasted with a cylinder (Expt. 1), the sphere was rated more rigid than the cylinder by all subjects, and more coherent for three of the four subjects. The higher rigidity ratings for spheres overall was actually due to a strong interaction between form and perspective: rigidity ratings were differentially lower for cylinders under perspective. The sphere gives less representation to dots which are substantially affected by the projection factor (far from the axis of rotation), and the increase in perceived non-rigidity may have resulted in object breakdown (segmentation, incoherence) in some cases. Braunstein & Andersen (1984) also compared spheres to cylinders. They used each form as the base for elliptical distortions, and found that sensitivity to minor axis variation (flatness of the elliptical orbits) was slightly greater when the base form was a cylinder than when the base form was a sphere. However, this was a cross-experiment comparison with different groups of subjects in the different conditions.

Distribution. The effect of dot distribution (in the volume or on the surface) was generally small with significant individual variation (see Figs. 3 and 4 and Table 2). Close examination of Table 2 suggests that distribution is more important when numerosity is large and when there is reduced single-view shape information (i.e., for the cylinder, Expts. 2 and 3). In these conditions, distribution in the volume increases depth ratings for three of the four subjects, and may improve coherence for some subjects as well.

Green's (1961) overall "goodness" measure showed slightly higher scores for surface representations than for completely random placements in the volume of a cube, but an enormous benefit for point representations with regular placements in the volume. Our random sampling incorporated partition-equality and hence represents a compromise between Green's random and regular conditions. Dots in the volume may increase depth ratings because a range of intermediate velocities is represented in the cylinders, while in surface representations of cylinders, all dots are traveling at more nearly the same velocity, except at the edges of the object. To the extent that differential velocity supports depth segregation (Braunstein & Andersen, 1981) representation of intermediate velocities may be useful. The ability of distribution to strongly effect the kinetic depth percept may also depend on the unavailability of other strong cues to shape such as perspective, texture density, or contour (see the discussion of Figure 5 below).

Perspective. For all subjects, adding polar perspective decreased rigidity ratings. The effect of perspective on coherence was small and depended on the subject. A collateral

analysis of the polar perspective trials that sorted those occasions on which the perceived rotation direction disagreed with the intended rotation direction found that most, but not all, of the decrease in rated rigidity with polar projection occurred when the observer perceived the stimulus in the reversed mode (see also Gregory, 1970; Schwartz & Sperling, 1983). Thus, when polar displays are perceived in their reversed mode, they appear grossly nonrigid; when polar displays are perceived veridically, they appear slightly less rigid than the corresponding parallel displays.

Neither our polar nor our parallel stimuli were viewed at the appropriate viewing distance. The parallel stimuli would have to be viewed from infinity; the polar stimuli from 6 cm; the actual viewing distances were in the range 1 m in the various experiments. (Had we produced appropriately projected objects for the 1 m viewing distance, they would have been negligibly different from the actual parallel stimuli. When viewed at the appropriate viewing distance, our polar displays possess little depth – largely a consequence of the large scale.) The greater mismatch between appropriate viewing distance and the actual viewing distance for polar stimuli conceivably might account for veridically perceived polar stimuli receiving slightly lower rigidity ratings than parallel stimuli. But this distance mismatch does not bear on the overwhelming cause of nonrigidity in polar displays–perceiving stimuli in reversed mode. Even the secondary effect of polar projection on rated rigidity may depend only weakly on projection/viewing distance mismatch. As noted previously, a pilot study in which object size was varied by a factor of 2:1, (producing a change between projected and and actual viewing distance of 2:1) had no significant effect on any rating. Similarly, Farrell (personal

communication) found subjects unable to accurately match simulated to actual viewing distance, holding object size constant. Finally, Cutting (1987) finds little impact of mismatch between simulated and actual viewing distances.

These rigidity and coherence results agree with the reported relation between amount of perspective and the ability to infer the intended rotation direction (Braunstein, 1977; Petersik, 1979). The difference in the effect of perspective on rated rigidity and on rated coherence may explain the inconsistent results of Braunstein, who found that perspective decreased a combined rating of coherence and rigidity in one study (1962), but had no effect in another (1977).

Perspective generally increased the rated depth (shape) of the percept (although not all contrasts were significant, see Table 2). A prior study by Braunstein (1962, see Table 3) also found that perspective improved a "strength of depth" rating. Petersik (1980) found that depth judgments were not affected by perspective. However, this same study found no effect of numerosity (N=5-60) on depth, which suggests that the experiment had insufficient power.

Interaction of perspective × distribution. Adding polar perspective increases rated depth, as does distributing dots throughout the volume of the object. Combining the two does not lead to a further increase. This interaction between perspective and dot distribution is illustrated in Figure 5, and significance levels for each subject are listed in Table 2. (Here again, the effect of distribution was greater in high-numerosity cylinders, Expt. 2 & 3.) As suggested above, some factors such as distribution may be more likely to control the percept in

the absence of other strong cues to shape.

Figure 4 & 5 about here.

A large individual difference. Occasionally, individual differences were very striking. An example of this is shown in Fig. 4. Here the coherence ratings for all conditions of Expt. 2 are shown for individual subjects. Subject RHS is the only subject for whom increasing the number of dots and distributing dots in the volume of the object decreases the coherence of the kinetic depth percept. Individual differences presumably occurred in earlier studies, but were undetected because prior studies collected few observations from each subject, and performed cross-subject analyses.

Three ratings are more informative than one. So far, we have been describing the empirical results with respect to manipulations of dot numerosity, perspective, and so forth. What is perhaps most important is the added information gained by having multiple ratings of the stimuli. These ratings, each of which can be (and has been in the literature) construed as a measure of the "strength" of a KDE percept, do not necessarily covary. In many cases, as we have seen, an experimental manipulation has a different effect on different ratings. For example, shape significantly affects mean ratings of rigidity, but has little or no effect on depth ratings. At high numerosity, further increases in numerosity continue to increase depth ratings but do not affects coherence ratings, and so forth.

Correlations between ratings. It is also possible to do a finer-grained analysis of the three ratings which looks beyond the means to the correlations on a trial by trial basis. Table 4 gives the trial-by-trial correlations pooled over conditions and subjects. Seven of the nine correlations in Table 4, are between -0.12 and +0.19; and the two highest correlations are still rather low, each relation accounting for only about 12% of the variance. Although the subjects are affected somewhat differently by some of the experimental factors, on the whole, different subjects tend to use the ratings quite similarly. Such low correlations between ratings clearly rule out a simple, single factor interpretation which would require high positive or high negative correlations between each pair of ratings. For every subject, at least two of the three inter-rating correlations are low. Obviously, the rated qualities of kinetic depth percept reflect at least two underlying dimensions. While experiments were not designed in a way that would expose the KDE depth percept to multidimensional analysis, they are sufficient to bring this inherent multidimensionality to the fore. That different ratings weigh differently on these dimensions cannot continue to be overlooked in KDE research.

Table 4 about here.

Discussion

Wide range of percepts. The KDE for a multi-dot display is quite rich. When viewing a stimulus with a small number of dots, there generally are many possible stable percepts. Even though the whole was geometrically derived from a rigid object, perceptually, subgroups of dots form clusters, and each subgroup appears to move independently in 3D, acting as a separate object. Groups of 2 or 3 dots can be perceived as moving independently in the plane, or as a 3-dimensional and rigid configuration, or as a non-rigid 3D configuration similar to the Ames window (as in Gillam, 1975, 1976, where lines segments were used rather than dots). In short, group of dots do not necessarily cohere as single objects, even when they are being perceived in depth and when a unitary rigid 3D interpretation is available. Even when dots are correctly perceived as parts of one coherent object, there is a range of possible percepts that differ in shape, in depth, and in perceived rigidity. The perceived coherence, shape and rigidity have a complex and partially decoupled relationship.

Decoupled aspects of percepts. Some degree of decoupling between aspects of a KDE percept has been known since Braunstein (1962) varied perspective in KDE displays and observed an inverse effect on mean judgments of *depth* and of combined *rigidity/coherence.* However, it has implicitly been assumed that the manipulation of rigidity by perspective was a special case. The current experiments demonstrate that this de-coupling between the various aspects of the percept is not restricted to the independent variation of perceived rigidity but is quite general because different factors affect different judgments. In terms of mean ratings in

our experiments, rated depth was significantly affected by numerosity, distribution and perspective. Rated segmentation was affected primarily by numerosity; this effect reflected a division between sparse and dense levels of numerosity (above or below 16 elements). Secondarily it was affected by form and perspective. Rated rigidity was primarily affected by perspective, and numerosity. Additionally, correlations between the three ratings were low and sometimes negative when measured on a trial-by-trial basis.

Experiments in the literature on multi-dot KDE have used as the dependent measure either ratings or paired comparisons on some judgment dimension (see Table 3). The judgment dimensions either selected from a variant of the three ratings used here, or combined two or more in one rating. Conflation of the dependent measure may, in part, explain some of the inconsistencies in the literature noted above. In particular, the combined coherence and rigidity ratings of Braunstein (1962, 1977) may account for the inconsistent effect of perspective in those studies.

Importance of independent factors. Our experiment manipulated a number of factors within a subject, factors that previously had been examined in separate experiments, or had been chosen arbitrarily as fixed factors which happened to differ, along with the dependent measure, between studies in the literature. Shape, distribution and numerosity have usually varied haphazardly between experiments. For example, Braunstein (1962, 1977) found inconsistent patterns of perspective on a rigidity-coherence judgment using 2-6 dots in the volume of a cube and 1000 dots in the volume of a sphere, respectively. Braunstein (1962)

and Petersik (1980) found inconsistent patterns of perspective on depth judgments using 2-6 dots in the volume of a cube and 4-45 dots in the volume of a sphere, respectively. It has been difficult to know whether the structural and numerosity factors explained the inconsistency in patterns.

The results of our experiments can be viewed as filling in Table 3 with a self-consistent set of data, and providing previously unavailable data in the empty cells. The results clearly separate three important aspects of a kinetic depth percept: depth (shape), coherence, and rigidity. Because our stimulus parameters are manipulated within subject, cells are directly comparable. A number of our results are similar to Green (1961), Braunstein (1962, 1977), and others. In other cases, where inconsistent findings were reported, we *suggested* explanations based on confounded dependent measures. Typically, inconsistent results between KDE experiments result from judgments that combine component aspects (e.g., depth, coherence, rigidity) in unspecified, but probably different weightings.

Nonmotion cues to depth in KDE displays. There are two classes of cues to object structure in our displays: static cues such as density and 2D object contour, and dynamic or motion cues which depend either on optic flow or on changing interpoint distances. Based on our data, the greatest likelihood of perceiving the veridical shape occurs with perspective images of spheres (rather than cylinders), with high dot numerosity, and with dots in the volume (rather than on the surface).

High dot numerosity guarantees a good representation of the 2D contour of the sphere and of 2D density cues. Even when the 2D contour is not as suggestive, as in the case of the cylinder, the density cues which become visible with high numerosity may be important in providing static cues to shape. (In the case of surface distribution of elements, the 2D density will increase toward the edges, while in volume distribution of elements the density cue is reversed, as in Fig. 1.) The presence of 2D cues to shape, whether from contour or density, may constrain the perception. High element numerosity also minimizes the likelihood of atypical clumping or grouping characteristics likely in low numerosity figures, which then are likely to cause grouped or segmented (i.e., incoherent) percepts.

Perspective may simply serve as an additional cue to depth organization. Alternatively, the exaggerated perspective used here may be especially effective because it slightly increases the proportion of elements moving in the same direction, yielding a display similar to that arising from an image with occlusion, and possibly allowing stronger input to an optic flow analysis at high numerosity (suggestion due to J. Todd, 1987).

Distribution in the volume provides a range of velocities in any local area, and may support a full depth percept by relating distance from the axis of rotation to dot velocity. Dot fields of different velocity, whether adjacent or superimposed, tend to segregate in depth (Braunstein & Andersen, 1981).

Percept description versus objective task measures. An alternative to the measurement of one or another aspect of the kinetic depth percept by rating is to conceptualize a different sort

of question. In rating, we ask about various aspects of the percept itself. An alternative is to ask whether a percept, whatever its subjective appearance, is adequate to support objective performance on a particular kind of judgment, such as a judgment of shape. Some attempts have been made in this regard. For example, Todd (1984) required subjects to make objective curvature judgments under various levels of non-rigidity in the kinetic depth image. Lappin et. al. (1980) required subjects to make objective (paired comparison) judgments of the degree of correspondence in two-frame displays. We have investigated one possible objective measure of having perceived shape from a kinetic depth display (Landy, Sperling, Dosher & Perkins, 1987; Sperling, Landy, Dosher & Perkins, 1987). This objective measure requires subjects to identify the object perceived from among a large lexicon of possible objects, and offers an attractive alternative to the elaboration of subjective methods under study here.

Relation to models. Three classes of computational models have been proposed to account for the kinetic depth effect based on motion cues: those deriving shape from optic flow fields (Clocksin, 1980; Koenderink & van Doorn, 1986), those deriving analytic solutions assuming rigidity from m views of n points (Hoffman & Bennett, 1985; Ullman, 1979), and those based on maximizing rigidity in interpoint distances (Ullman, 1979, 1984; Hildreth & Grzywacz, 1986; Landy, 1987). Usually, flow-field models are applied to objects composed of densely packed points, and interpoint-distance models are applied to images composed of less than a few dozen points. Interpoint-distance models apply geometric computations to the 2D image-plane positions of the given points to compute a 3D object that is either totally rigid (Ullman, 1979) or a 3D object that deforms minimally between adjacent frames (Landy, 1987;

Ullman, 1984).

We will illustrate the problems that rigidity models have with data such as ours by considering, as an example, the incremental rigidity algorithm of Ullman (1984). When an n-point 3D object undergoes rotation, the algorithm takes as its input a sequence of frames that represent the 2D image-plane x,y projections of the n points. For each frame, the algorithm outputs an estimated depth value z for each point plus one overall fidelity score. The computation consists of a gradient descent in the space of depth values z to maximize the fidelity criterion. This criterion measures the deformation (nonrigidity) in the recovered 3D object between the current frame and the prior frame.

To evaluate such an algorithm as a *psychological* model, one must associate quantities produced by the algorithm with aspects of human perception. We have shown here three aspects of performance which are partially separable in performance measures: segmentation, depth and rigidity. Consider what happens in the case of 4-point displays. Ullman's algorithm, like most others, simply assumes element correspondence and figural segmentation as prior processes. For four-point objects, the incremental rigidity algorithm would have recovered the veridical single object with rigid depth assignments for all nonperspective images in the experiment. On the other hand, only one of our four subjects regularly perceived 4-point displays as unitary; for the other subjects, these were usually perceived as two or more objects moving independently. This grossly violates the segmentation assumed by the Ullman model.

The algorithm's estimated depth values seem a plausible basis for predicting human depth judgments, and the algorithm's fidelity score seems a plausible basis of rigidity judgments. An immediate problem is that perspective-dependent modulations in position of elements on the image plane are treated as noise by this (and most other) algorithms (Sperling & Dosher, 1987), while our subjects' depth percepts are improved by moderate amounts of perspective. The problems surrounding predictions with parallel and with perspective projections are particularly enlightening. Detailed consideration (Sperling & Dosher, 1987; Dosher & Sperling, 1988) showed that a class of models including Ullman (1984) exhibit unrealistic properties because they do not incorporate any perspective transformation, whereas the models would require a flexible perspective transformation to deal with perceptual facts. Parallel perspective algorithms, such as Ullman's, when applied to perspective images, such as those in our experiments, yield flattened depth estimates relative to nonperspective images. On the contrary, for our human observers, perspective increased depth ratings slightly.

Predictions of perceived rigidity based on the fidelity criterion are the most problematic aspect of Ullman's algorithm. When the image is produced by perspective transformation, parallel-perspective rigidity algorithms (e.g., Ullman) cannot distinguish between veridical and reversed depth 3D recovered objects – these yield precisely equal rigid and nonrigid solutions. For our subjects, reversed depth perceptions are grossly more nonrigid than the veridical ones, a powerful perceptual fact that is beyond the scope of this class of models.

Purely geometric algorithms that yield explicit solutions to three dimensional objects given m-views of n-points (Bennett & Hoffman, 1985; Hoffman & Bennett, 1985; Ullman, 1979; Webb & Aggarwal, 1981) fare much worse than the incremental rigidity algorithm. Again, segmentation is simply assumed. The algorithms yield exact solutions under certain conditions when the stimuli represent rigid objects. The outputs here are exact solutions, or the fact that a solution failed. An exact solution must be rigid, so the model cannot predict any particular nonrigid percept, nor does it have computational by-products which could support rigidity-nonrigidity judgments, partial depth, or incomplete segmentation.

We conclude that, while many existing algorithms are of great interest as a possible basis *for robotics solutions* to the structure from motion problem, they are inadequate as psychological models. Our experiments suggest that a successful psychological model must identify at least three separable aspects of recovered objects that could serve as a basis for the three separable, measurable aspects of kinetic depth perception.

ACKNOWLEDGMENT

The work described in this report was carried out in the Psychology Department of New York University and supported by Office of Naval Research, Grant N00014-85-K-0077 and USAF Life Sciences Directorate, Visual Information Processing Program Grants 85-0364 and 88-0140.

REFERENCES

Andersen, G. J., and Braunstein, M. L. (1983). Dynamic occlusion in the perception of rotation in depth. *Perception & Psychophysics*, **34**, 356-362.

Bennett, B. M., and Hoffman, D. D. (1985). The computation of structure from fixed-axis motion: Nonrigid structures. *Biological Cybernetics*, **51**, 293-300.

Braunstein, M. L. (1962). Depth perception in rotating dot patterns: Effects of numerosity and perspective. *Journal of Experimental Psychology*, **64**, 415-420.

Braunstein, M. L. (1977). Perceived direction of rotation of simulated three-dimensional patterns. *Perception and Psychophysics*, **21**, 553-557.

Braunstein, M. L., and Andersen, G. J. (1981). Velocity gradients and relative depth perception. *Perception and Psychophysics*, **29**, 145-155.

Braunstein, M. L., and Andersen, G. J. (1984). Shape and depth perception from parallel projections of three dimensional motion. *Journal of Experimental Psychology: Human Perception and Performance*, **10**, 749-760.

Caelli, T. M. (1979). On the perception of some geometric properties of rotating three dimensional objects. *Biological Cybernetics* **33**, 29-37.

Caelli, T. M. (1980). Amplitude, frequency and phase determinants of perceived rotations and rigidity in the kinetic depth effect. *Biological Cybernetics* **36**, 213-219.

Clocksin, W. F. (1980). Perception of surface slant and edge labels from optical flow: a computational approach. *Perception* **9**, 253-269.

Cutting, J. E. (1987). Rigidity in cinema seen from the front row, side aisle. *Journal of Experimental Psychology: Human Perception and Performance*, **13**, 323-334.

Dosher, B. A., Sperling, G., and Wurst, S. A. (1986). Tradeoffs between stereopsis and proximity luminance covariance. *Vision Research*, **26**, 973-990.

Dosher, B. & Sperling. G. (1988). Predicting rigid and nonrigid perceptions. Manuscript.

Dosher, B. A., Landy, M. S., & Sperling, G. (1988). Kinetic depth effect and optic flow. I. 3D Shape from Fourier motion. *Mathematical Studies in Perception and Cognition*, NYU Memorandum **88-3**.

Gillam, B. (1975). New evidence for "closure" in perception. *Perception and Psychophysics*, **17**, 521-524.

Gillam, B. (1976). Grouping of multiple ambiguous contours: towards an understanding of surface perception. *Perception*, **5**, 203-209.

Green, B. F., Jr. (1961). Figure coherence in the kinetic depth effect. *Journal of Experimental Psychology*, **62**, 272-282.

Gregory, R. L. (1970). *The Intelligent Eye*. New York: McGraw-Hill, 1970.

Hildreth, E. C., and Grzywacz, N. M. (1986). The incremental recovery of structure from motion: Position vs. velocity based formulations. Proceedings of the Workshop on Motion: Representation and Analysis, IEEE Computer Society #696, Charleston, South Carolina, May 7-9, 1986.

Hoffman, D. D., and Bennett, B. M. (1985). Inferring the relative three-dimensional positions of two moving points. *Journal of the Optical Society of America A*, **2**, 350-355.

Koenderink, J. J., and van Doorn, A. J. (1986). Depth and shape from differential perspective in the presence of bending deformations. *Journal of the Optical Society of America A*, **3**, 242-249.

Kropfl, W. J. (1975). Variable raster and vector display processor. *Unpublished Technical Memorandum TM-75-1223-3*, Bell Telephone Laboratories, 1975.

Landy, M. S. (1987). A parallel model of the kinetic depth effect using local computations. *Journal of the Optical Society of America A*, **4**, 864-877.

Landy, M. S., Dosher, B. A., and Sperling, G. (1986). Assessing kinetic depth in multi-dot displays. *Bulletin of the Psychonomic Society*, **19**, 23.

Landy, M. S., Sperling, G., Dosher, B. A., and Perkins, M. E. (1987). Structure from what kinds of motion? *Investigative Ophthalmology and Visual Science, ARVO Supplement*, **28**, 233.

Lappin, J. S., Doner, J. F., and Kottas, B. L. (1980) Minimal conditions for the visual detection of structure and motion in three dimensions. *Science*, **209**, 717-719.

Petersik, J. T. (1979). Three-dimensional object constancy: Coherence of a simulated rotating sphere in noise. *Perception & Psychophysics*, **25**, 328-335.

Petersik, J. T. (1980). The effects of spatial and temporal factors on the perception of stroboscopic rotation simulations. *Perception*, **9**, 271-283.

Schwartz, B., & Sperling, G. (1983). Luminance controls the perceived 3-D structure of dynamic 2-D displays. *Bulletin of the Psychonomic Society*, **21**, 456-458.

Sperling, G., & Dosher, B. A. (1987). Predicting rigid and nonrigid perceptions. *Investigative Ophthalmology and Visual Science, ARVO Supplement*, **28, No. 3**, 362.

Sperling, G., Landy, M. S., Dosher, B. A., and Perkins, M. E. (1987). The kinetic depth effect and identification of shape. Mathematical Studies in Perception and Cognition 87-1, New York University.

Todd, J. T. (1984). The perception of three-dimensional structure from rigid and non-rigid motion. *Perception and Psychophysics*, **36**, 97-103.

Todd, J. (1987). The kinetic depth effect and apparent motion. Talk at New York University, March.

Ullman, S. (1979). *The Interpretation of Visual Motion.* Cambridge, MA: MIT Press.

Ullman, S. (1984). Maximizing rigidity: the incremental recovery of 3-D structure from rigid and non-rigid motion. *Perception*, **13**, 255-274.

Wallach, H., & O'Connell, D. N. (1953). The kinetic depth effect. *Journal of Experimental Psychology*, **45**, 205-217.

Webb, J. A., & Aggarwal, J. K. (1981). Visually interpreting the motion of objects in space. *Computer*, **14**, 40-49.

FIGURE LEGENDS

Figure 1. Single frames from some sample stimuli varying in numerosity, distribution and form.

Figure 2. The inferred "top-views" of the stimuli that were used to define the 5 levels of perceived depth ratings.

Figure 3. Response data for all ratings and stimulus manipulations in Experiment 1 (a) and 2 (b), pooled across subjects. The parameter is the particular rating made: perceived depth (circles), rigidity (squares), and coherence (triangles). In each panel, the first set of curves is for the number of dots in the stimulus, the second for the effect of distributing the dots across the surface or throughout the volume of the object, and the third for the effect of perspective transformation (parallel or polar).

Figure 4. Coherence ratings in Experiment 1. The parameter is the particular subject. Note the large individual differences.

Figure 5. Interaction between perspective transformation and dot distribution in Experiment 1 and 2 (pooled).

Table 1.

Experimental Factors and Conditions

	Numerosity	Form [cylinder sphere]	Distribution [surface volume]	Perspective [parallel polar]	Luminance
Exp. 1	4,8,16,36	both	yes	yes	1.45 μcd/dot(a)
Exp. 2	36,48,64,80	cylinders	yes	yes	3.02 μcd/dot(b)
Exp. 3	48	cylinders	yes	yes	0.86,3.02 11.52,42.72 μcd/dot(b)

Note:

(a) Point plot display, resolution 1024 x 1024; 36 new frames per 360 *deg* rotation (or 10 *deg* per frame); 60 msec per new frame, or 2.16 sec per full rotation; dark room; binocular free viewing; viewing distance 1.1 meters; object diameters 2 *deg* visual angle (parallel perspective).

(b) Raster display, resolution 512 x 512; 36 new frames per 360 *deg* rotation (or 10 *deg* per frame); 66.67 msec per new frame, or 2.4 sec per full rotation; dim room (8 cd/m^2) with light-tight viewing hood; monocular viewing through a reduction aperture; viewing distance 1.6 meters; object diameters 2 *deg* visual angle (parallel perspective).

Table 2.

Significant Factors in Experiments

Depth Judgment

	Numerosity	Form	Distribution	Perspective	D X P
Exp. 1	[small #s]				
ML	+ ***	ns	+ ***	ns	**
BD	+ ***	ns	+ ***	+ **	*
CS	+ ***	ns	+ ns	ns	ns
RS	U ***	+ **	- **	ns	ns
Exp. 2	[large #s]				
ML	~ **	----	+ ***	+ ***	***
BD	~ **	----	+ ***	+ **	*
CS	~ *	----	+ **	ns	†
RS	~ ***	----	- ***	ns	ns
Exp. 3					
ML	----	----	+ **	+ ***	†
BD	----	----	+ *	+ *	*
CS	----	----	+ ns	+ *	†
RS	----	----	- *	ns	ns

Table 2. (continued)

Significant Factors in Experiments

Segmentation Judgment

	Numerosity	Form	Distribution	Perspective	D X P
Exp. 1	[small #s]				
ML	+ ***	+ *	ns	- ***	***
BD	+ ***	ns	+ **	+ **	*
CS	+ ***	+ **	ns	- ***	ns
RS	+ ***	+ **	- ***	- **	ns
Exp. 2	[large #s]				
ML	ns	----	ns	ns	ns
BD	ns	----	+ ***	+ ***	***
CS	ns	----	- **	- ***	ns
RS	ns	----	- ***	- ***	ns
Exp. 3					
ML	----	----	+ *	ns	†
BD	----	----	+ **	+ *	**
CS	----	----	+ *	- *	ns
RS	----	----	- **	- *	*

Table 2. (continued)

Significant Factors in Experiments

Rigidity Judgment

	Numerosity	Form	Distribution	Perspective	D X P
Exp. 1	[small #s]				
ML	+ ***	+ **	+ †	- ***	**
BD	+ ***	+ **	+ **	- ***	ns
CS	+ **	+ **	+ †	- ***	ns
RS	X ***	+ *	+ *	- ***	*
Exp. 2	[large #s]				
ML	ns	----	ns	- ***	ns
BD	~ **	----	+ **	- ***	ns
CS	~ *	----	- **	- ***	ns
RS	ns	----	- ***	- ***	ns
Exp. 3					
ML	----	----	ns	- ***	ns
BD	----	----	+ *	- ***	ns
CS	----	----	ns	- ***	ns
RS	----	----	ns	- **	ns

Table 2. (continued)

Legend:

Note: The † indicates $p<.10$, * indicates $p<.05$, ** indicates $p<.01$, and *** indicates $p<.001$. These p-values are the significance of corresponding F-values from an analysis of variance for each subject treating rotation direction and tokens of stimuli as the random factor. The symbols ˜, + and - indicate the pattern of the effect, and can be referenced to the legend list below for each factor. See the text for a discussion of the interaction of distribution and perspective.

Numerosity: + increasing with # points

˜ saturates with large # points

U u-shaped function of # points

Form: + sphere > cylinder

Distribution: + volume > surface

Perspective: + polar > parallel

Table 3.

Summary of Related Results

	Numerosity	Form [cylinder sphere]	Distribution [surface volume]	Perspective [parallel polar]
Depth	=Petersik,1980			+Braunstein,1962 =Petersik,1980
Segment	-Braunstein,1962			
Rigidity	=Petersik,1980			-Petersik,1979 -Braunstein,1977
Combined	+~Green,1961	-Braunstein&Anderson,1984	-Green,1961	-Braunstein,1962 =Braunstein,1977

Table 3(continued)

Legend:

Note: The symbols ~, + and - indicate the pattern of the effect, and can be referenced to the Note below Table 2. A summary of some relevant factors in prior experiments follows:

Braunstein (1962). Dots in volume of cube. N=2-6. Depth and coherence/rigidity judgments.

Braunstein (1977). Dots in volume of sphere. N=1000. Varied perspective in horizontal and vertical dimensions. Direction and coherence/rigidity judgments.

Braunstein & Anderson (1984). Dots on surface of sphere or elipses. N=140-160. Shape and quality judgment.

Green (1961). Dots or line elements in volume or on surface of cube. N=4-64. *Goodness rating* (combined segmentation and rigidity).

Petersik (1979). Dots in volume of sphere. N=4-45. Depth and direction rating.

Petersik (1980). Dots in volume of sphere. N=5-60. Depth and direction rating.

Table 4.

Correlations Between Judgments: All Subjects

	Expt. 1	Expt. 2	Expt. 3
Depth-Rigidity	.05	-.12	-.01
Depth-Segmentation	.08	.07	.16
Rigidity-Segmentation	.36	.35	.19

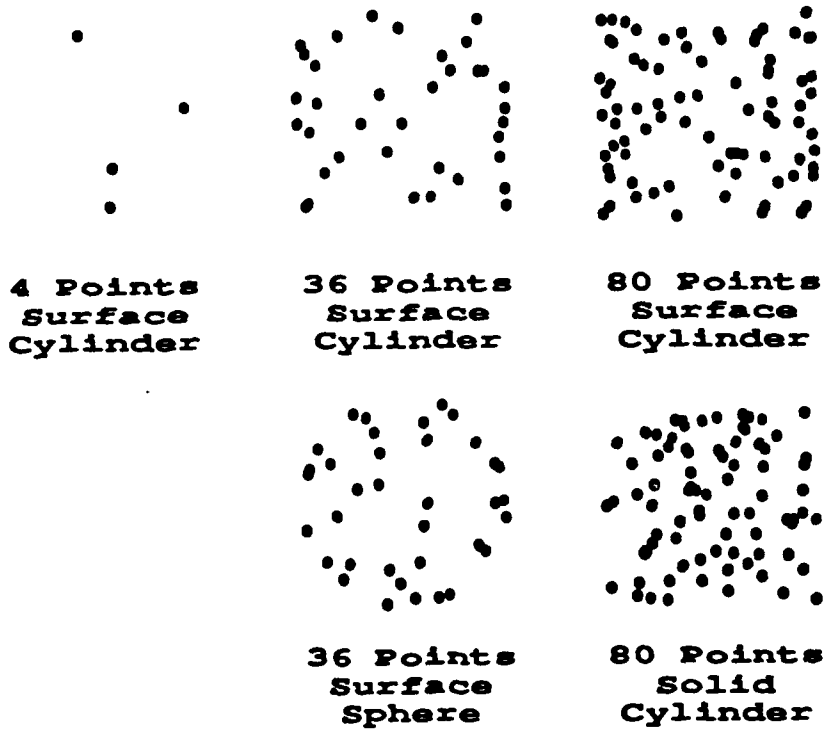

Figure 1. Single frames from some sample stimuli varying in numerosity, distribution and form.

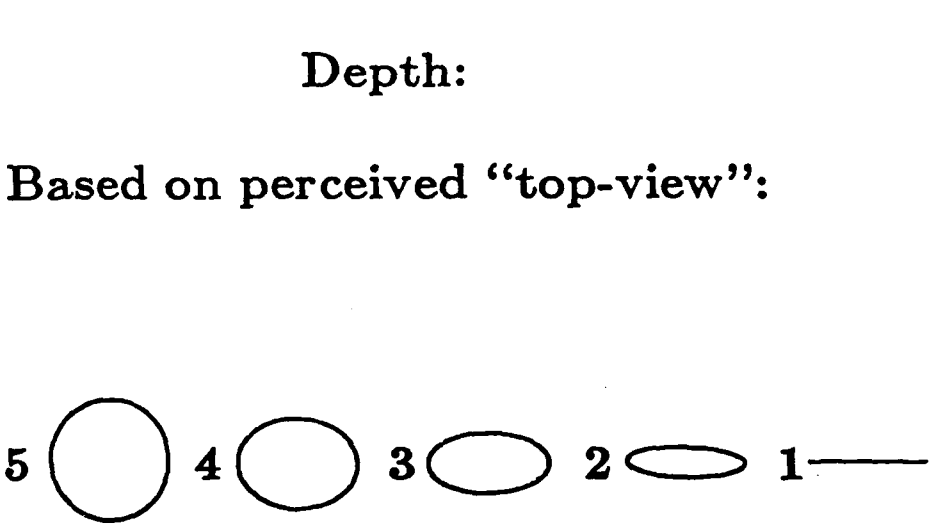

Figure 2. The inferred "top-views" of the stimuli that were used to define the 5 levels of perceived depth ratings.

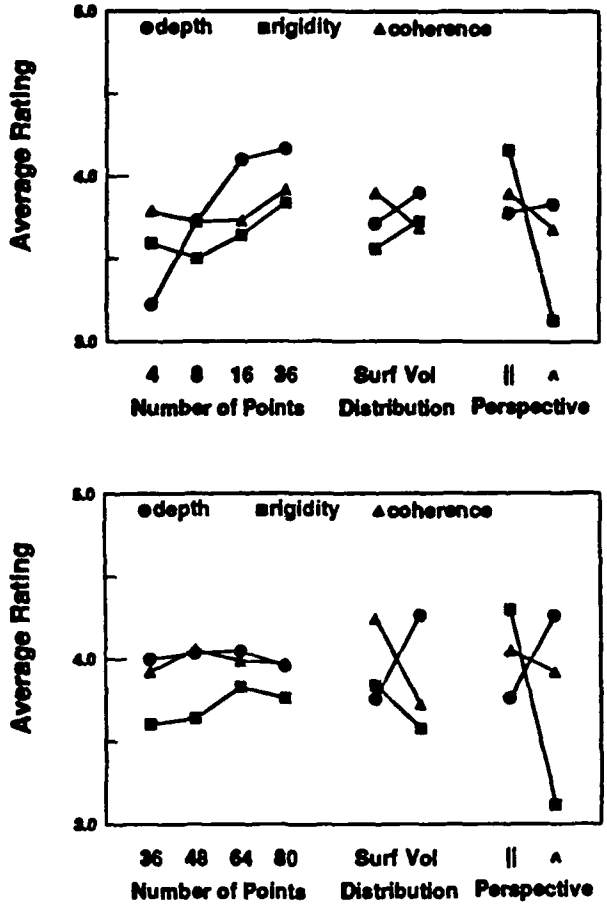

Figure 3. Response data for all ratings and stimulus manipulations in Experiment 1 (a) and 2 (b), pooled across subjects. The parameter is the particular rating made: perceived depth (circles), rigidity (squares), and coherence (triangles). In each panel, the first set of curves is for the number of dots in the stimulus, the second for the effect of distributing the dots across the surface or throughout the volume of the object, and the third for the effect of perspective transformation (parallel or polar).

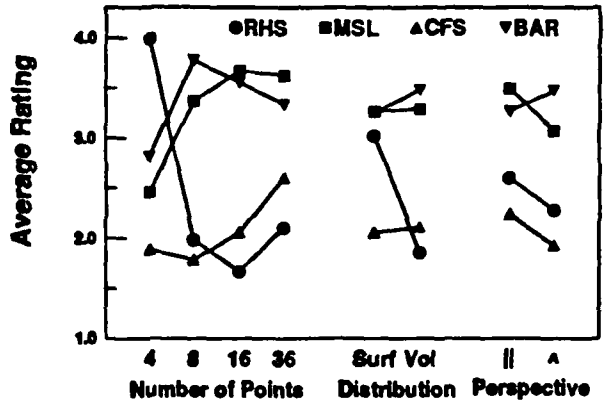

Figure 4. Coherence ratings in Experiment 1. The parameter is the particular subject. Note the large individual differences.

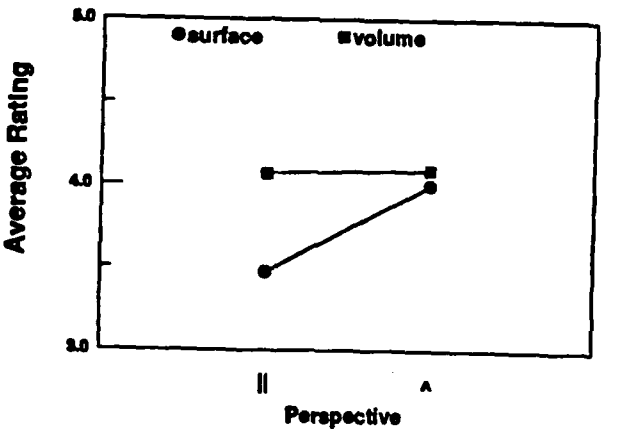

Figure 5. Interaction between perspective transformation and dot distribution in Experiment 1 and 2 (pooled).